Robust Portfolio Optimization based on Evidence Theory

Amirhossein Eskoruchi¹, Emran Mohammadi¹*, Seyed Jafar Sajadi¹
¹School of Industrial Engineering, Iran University of Science and Technology, Tehran 16846-13114, Iran

Abstract
During the past few years, there have been some turbulent events on world' economy, which have significantly influenced the performance of companies. Therefore, there is an urgent need to use a robust method to handle the existing uncertainties on the performance of the companies. This paper uses Evidence Theory to present an innovative and practical approach to consider the experts’ opinions which are based on the available evidence regarding the factors that influence the stock market. Subsequently, the study proposes a way to determine the changes in these factors from possible scenarios on historical data to find the return range of different investment alters to be used in robust optimization. Moreover, in a case study, the sensitivity of the Iranian stock market to exchange rate fluctuations is examined under a set of scenarios which was due to the ambiguousness of a unified viewpoint of that rate’s value midst experts as one of the noticed factors. The preliminary results of a real-life case study reveal that the proposed approach is useful practically and productively.

Keywords
Portfolio Optimization; Evidence Theory; Robust Optimization; Hybrid Uncertainty; Fundamental analysis.

* Corresponding author. Tel.: +989370449466
E-mail address: e_mohammadi@iust.ac.ir
1. Introduction

Portfolio optimization is one of the most critical subjects in financial issues, where the influential work of Markowitz (1952) provided the fundamental portfolio model that is used as the foundation for modern portfolio theory. However, the valuable function of this framework in the finance administration industry has been discouraging. The fundamental drawback of Markowitz's model is that the idea assumes the appraisals of each stock returns and the variances are accurate but also tends. This issue is analytically reported and empirically tested by many researchers (see[1], [2], [3]and [4]). While the capital market is affected directly or indirectly by numerous factors, making estimation of assets' value is uncertain. Thus, detecting behavioral scenarios and prioritizing influential factors in terms of practical aspects help us overcome some unwanted uncertainties. In addition, in cases where there are historical data, although the data can be used, they may not necessarily end up with an efficient outcome. The stock market is influenced by economic, political, and social elements. In this regard, using the opinion of experts for the process of decision making can be very useful. The present study uses Evidence Theory (ET) in a new manner in portfolio robust optimization problems. The expenses incurred and development time has been reduced considerably compared with other existing models due to repeated expert's interactions as well as the overall complexity of the model which was the chief problem in most of the recent researches. At the same time, we incorporate another level of uncertainty handling mechanism into the portfolio selection model.

One of the most popular methods which can be utilized to deal with uncertainty is the Robust Optimization (RO) methodology [5]. In this regard, a novel robust portfolio selection and optimization approach is provided to deal with the vagueness of the data, boosting the robustness of investment procedure against uncertainty, declining computational complexity, and comprehensive assessments of stocks from various financial aspects and criteria are submitted. For this purpose, the outputs of the ET are been used as the inputs of the robust portfolio optimization model (RPOM) to create more stability in the efficiency of assets and to achieve the nominal portfolio optimization model in real-world conditions even when the empirical distribution of samples is deviated from the normal distribution. This paper is dealing with a hybrid uncertainty in portfolio optimization. After identifying investment options and determining the factors affecting stock returns, we first rank the stocks using the data envelopment analysis approach, filter them based on maximum return and minimum risk, and determine how stock prices are affected by influential factors according to the Sharp multi-factor model. In the portfolio optimization literature, many studies have been conducted to improve the performance of portfolio optimization different nominal models under various circumstances (see[6], [7], [8], [9], [10], [11], [12],[13]).

Furthermore, the current strategy in the investment portfolio might change when an unexpected event or an incident alters an investor's environmental situations. Such changes require a logical and comprehensive evaluation of the portfolio for striking a tradeoff [14]. In this case, various investigations have been conducted based on robust optimization (RO) to face
uncertain conditions and parameters related to a portfolio. Ben-tal et al. [15] modeled a robust multi-objective portfolio selection problem using a linear programming approach. According to Moon and Yao [16], a robust absolute value of deviation from the mean model is determined. Kim and Fabozzi [17] discuss the use of robust factor investing in portfolio management. Using a robust approach, Goldfarb and Iyengar [18] introduced a conic programming model for single-period portfolio selection problems by taking advantage of the factor model in portfolio efficiency. Ling and Xu [19] concentrated on giving option contracts in a portfolio using an ellipsoidal uncertainty set with a joint margin. In their method, the risk is controlled by an option contract and existing strategies. Carvalho et al. [20] find the limiting portfolios of RO, formulated with different uncertainty matrices, for the highest and lowest uncertainty levels. Quarant et al. [21] used the Ben-Tal and the Nemirovski's approach to develop a conditional value-at-risk robust model that was a non-linear model.

Bertsimas and Pachamanovab [22] provided a robust optimization for multi-period portfolio selection problems under a number of scenarios. In this case, subsequent portfolio efficiency is considered uncertain coefficients in the optimization problem and an investor's risk-aversion level as the fluctuation level of total efficiency estimation error. Chen and Tan [23] developed a nominal mean variance model using the stochastic optimization method and robust optimization. They assumed that the assets' efficiency was uncertain and used the Bertsimas and Sim approach for RO development. Ghahtarani and Najafi [24] developed a robust multi-objective portfolio selection model for which they used goal programming. Using ellipsoidal uncertainty sets, Pinar and Paç [25] developed their modeling for the single-period and multi-period mean-semi variance model. Ling et al. [26] investigates robust multi-period portfolio selection based on downside risk with asymmetrically distributed uncertainty.

Considering mean vector components and covariance matrix in distance regions, Tütüncü and Koenig [27] solved the portfolio selection problem to achieve a robust answer with non-linear programming. Unlike most researchers in this field, Bertsimas and Sim [28] examined optimization under an uncertain condition with models in which the parameters were not modeled by polygonal sets other than ellipsoidal sets. Afterward, Bertsimas and Theley [29] provided a model where polygonal uncertainty sets are used instead of ellipsoidal sets. Dimmock et al. [30] provided a different work in uncertainty modeling and created a problem in robust portfolio selection, where certainty explained the probability of asset efficiency distribution. Ji and Lejeune [31] study a class of distributionally robust optimization model with Sharpe, Sortino–Satchel, and Omega ratios. Sharma et al. [32] consider a robust formulation of their respective models under discrete distribution of returns since the true distribution of returns is unknown. In addition, they evaluated the relationships between risk, robustification, and portfolio efficiency. They concluded that as robustness increases, portfolio risk and efficiency declines, diversifying the portfolio. Peykani et al. presented a robust bi-objective model for portfolio selection, which is capable to be used under uncertainty of financial data and also a NSGA-II meta-heuristic algorithm to solve the suggested model of research due to the complexity of the proposed model [33]. Georgantas et al. [34] conducted a comprehensive
empirical evaluation of various robust models performance for popular risk criteria in robust portfolio optimization. They also provided a broad comparative analysis of their performance using the US market, where more efficiency was observed in this type of model than nominal models. Ghahtarani et al. [35] provided a comprehensive review of recent advances in robust portfolio selection problems and their extensions, from both operational research and financial perspectives. Incorporating future returns scenarios in the investment decision process and developing a conventional minimax regret criterion formulation, in multi-objective programming problems, Mohammadi et al. [36] verify the validity of the proposed approach through an empirical testing application on the top 75 companies of Tehran Stock Exchange Market. Peykani et al. [37] presented six robust data envelopment analysis (RDEA) models based on the most widely cited and popular classic data envelopment analysis models in the first phase and two robust portfolio optimization models including robust mean-semi variance-liquidity and robust mean-absolute deviation-liquidity in the second phase. Finally, Xidonas et al. [38] provides a categorized bibliographic review which has a broad scope, yet is limited in technical details.

Having utilize the Fuzzy Delphi method in the first stage for crucial factors identification, Thakur et al. [39] hierarchically arranged the stocks by main factors and historical data based on ET. Finally, Ant Colony Simulation for portfolio optimization was used. Skoruchi and Mohammadi [40] examined the high reaction of the Iranian stock market to the dollar value perturbation in the diverse scenarios because of the vagueness of a unified opinion of the amount of that rate among experts and obtained its rate utilizing ET. As seen, the robust model's uncertainty parameters are given without acceptable justification and no admissible mechanism is used to ascertain range. To the best of our knowledge, it is the first time that ET is used to determine this interval according to the available evidence.

The rest of the paper is organized as follows: after a general investigation of the research subject, the fundamental aspects of investment choices based on financial ratios will be analyzed. In section three, ET will be further introduced to receive experts' view under uncertain situations. In section four, we will review the Symbiotic Organisms Search algorithm to answer the problem numerically. Section five will explore how to conduct a portfolio utilizing actual datasets extracted from Tehran capital market. In section six, the obtained results will be provided as well as suggestions for further studies. Fig. 1 illustrates the conceptual flow chart.

(Insert Fig.1 almost here)

2. Fundamental analysis of investment alternatives based on DEA

Fundamental analysis is the process of evaluating a public firm for its investment worthiness by looking at its business at the basic or fundamental financial level, see for example, Thomsett[41]. Moreover, financial ratios are created with the use of numerical values taken from financial statements to gain meaningful information about a company. To begin with, liquidity ratios attempt to measure a company's ability to pay off its short-term debt obligations. This is
done by comparing a company's most liquid assets (or those that can be easily converted to cash) and its short-term liabilities. Profitability indicator ratios profitability Indicator Ratios much like the operational performance ratios, give users a good understanding of how well the company utilized its resources in generating profit and shareholder value. Debt ratios are the third series of ratios. A general idea of the company's overall debt load and its mix of equity and debt are given to users by these ratios. Debt ratios can be utilized to define the total level of financial risk an institute. As a result, the greater quantity of debt held by a company the greater risk of bankruptcy.

Operating performance ratios basically, determine how efficiently and effectively a firm is using its assets to generate sales and boost shareholder value. Generally, the better these ratios are, the better it is for shareholders. Cash flow indicator ratios indicators focus on the cash being generated in terms of how much is being generated and the safety net that it provides to the company. These ratios can give users another look at the financial health and performance of a company. Investment Valuation Ratios can be used by investors to estimate the attractiveness of a potential or existing investment and get an idea of its valuation.

3. Analysis of the effective factors based on Evidence Theory (Dempster–Shafer theory)

In the late 1960s, Dempster provided a new viewpoint regarding probability size in his famous paper Dempster [42]. He acknowledged that the classic structure of probability theory does not properly provide the possibility of demonstrating the lack of knowledge. However, in real-world conditions, decision-makers face a type of lack of knowledge in expressing their own subjective probability. In addition, objective probabilities can be aggregated, while empirical investigations do not demonstrate such a property regarding subjective probabilities, and it seems that these probabilities have sequential size. Despite providing a new concept, related researchers neglected this theory for a long time due to its weakness in explaining and establishing its own viewpoints in the mathematics form. Nevertheless, it was provided and redefined in the late 1970s in the form of a proper framework by a famous mathematician named Shafer [43]. One of the important instruments in defining uncertainty is ET, offering the opportunity for a decision-maker (DM) to realize new probabilities. This theory endures with the consultation with regard to existing beliefs of a situation or a system of situations. Beliefs of individuals are dissimilar when dealing with a single type of occurrence, though they can be inspected and combined by a particular method. Indeed, a number of beliefs caused by observation and perception of evidence has shaped Dempster–Shafer theory (DST).

In the following, DST for multi-criteria decision-making analysis with respect to uncertainty conditions will be briefly examine by Mohammadi and Makui [44]. Assume $X = \{x_1, x_2, \ldots, x_m\}$ is a set of options, $W = \{w_1, w_2, \ldots, w_n\}$ a set of weights, $A = \{a_1, a_2, \ldots, a_n\}$ a set of benchmarks, so that $0 \leq w_j \leq 1$, $1 \leq j \leq n$, $\sum_{j=1}^{n} w_j = 1$. Let assume $P$ is the evaluation rank of $H_1, H_2, \ldots, H_p$ for multi-criteria evaluation of options. Assume $\beta_{q,j}(x_i)$ indicates a belief degree of the fact that $a_j$ criterion has been assessed for $x_i$ with $H_q$ degree, $0 \leq \beta_{q,j}(x_i) \leq 1$ and
\[ \sum_{q=1}^{p} \beta_{q,j}(x_i) \leq 1. \] Assume \( S(a_j(x_i)) \) indicates criterion evaluation value \( a_j \) for \( x_i \) option, presented as follows.

\[ S(a_j(x_i)) = [H_q, \beta_{q,j}(x_i)] \tag{1} \]

Where \( H_q \) is an evaluation degree, so that \( 1 \leq q \leq p \), \( 1 \leq i \leq m \), \( 1 \leq j \leq n \). Criteria evaluation for options is defined in the form of a decision matrix \( D \), demonstrated as follows.

\[ D = (S(a_j(x_i)))_{nxm} \tag{2} \]

Where \( 1 \leq i \leq m \) and \( 1 \leq j \leq n \). Thereby, according to decision matrix \( D \), we can collect criteria evaluation value for \( x_i \) option as follows Mohammadi and Makui [44].

Initially, we convert belief degree \( \beta_{q,j}(x_i) \) about assessment degree \( H_q \) regarding \( a_j \) criterion of \( x_i \) option to basic probable mass \( m_{q,j}(x_i) \). Such that:

\[ m_{q,j}(x_i) = w_j \beta_{q,j}(x_i) \tag{3} \]

Where, \( 1 \leq q \leq p \), \( 1 \leq i \leq m \), \( 1 \leq j \leq n \).

Now, assume that \( m_{H,j}(x_i) \) indicates the probable residual mass of \( a_j \) criteria regarding the assessment of \( x_i \) option, which is defined as follows:

\[ m_{H,j}(x_i) = \bar{m}_{H,j}(x_i) + \tilde{m}_{H,j}(x_i) \tag{4} \]

\[ \bar{m}_{H,j}(x_i) = 1 - w_j \tag{5} \]

\[ \tilde{m}_{H,j}(x_i) = w_j \left( 1 - \sum_{q=1}^{p} \beta_{q,j}(x_i) \right) \tag{6} \]

Where, \( 1 \leq q \leq p \), \( 1 \leq i \leq m \), and \( 1 \leq j \leq n \). Probable residual mass that is not devoted to each assessment degree consists of two parts: The section pertaining to the violation in the assessment process and the part pertinent to relative weights of criteria [45].

\( \tilde{m}_{H,j}(x_i) \) is the first part of probable residual mass, which is not yet devoted to assessment degrees. According to the fact that \( a_j \) criterion plays a part in the assessment process according to its weight, that is \( w_j \), \( \tilde{m}_{H,j}(x_i) \) is a descending function of \( w_j \) [46]. \( \tilde{m}_{H,j}(x_i) \) will be equal to 1, if the weight of \( a_j \) is equal to zero or \( w_j = 0 \). \( \tilde{m}_{H,j}(x_i) \) will be zero if \( a_j \) dominates the assessment or \( w_j = 1 \). In other words, \( \tilde{m}_{H,j}(x_i) \) represents the degree to which other criteria could contribution in the assessment process.
The second part of the probable Hesitancy mass is $\tilde{m}_{H,j}(x_i)$, not yet devoted to an assessment degree, and due to violation in the assessment procedure of $S(a_j(x_i))$ will result. $\tilde{m}_{H,j}(x_i)$ will be zero if $S(a_j(x_i))$ is complete or $\sum_{q=1}^{p} \mu_{q,j}(x_i)=1$, otherwise $\tilde{m}_{H,j}(x_i)$ would be a positive value. $\tilde{m}_{H,j}(x_i)$ will be proportional to $w_j$, whose positive quantities will lead to the next constraint’s violation.

$G_{I(l)}$, a subset of $l$ number of first criteria, will be proposed as follows:

$$G_{I(l)} = \{a_1, a_2, \ldots, a_l\}$$

7)

Assume that $m_{q,I(l)}(x_i)$ is a probable mass that represents the support degree of a belief that all touchstone existing in $G_{I(l)}$ subset emphasize that $x_i$ option with $H_q$ degree is evaluated. $m_{H,I(l)}(x_i)$ presents the probable hesitancy mass that is not devoted to assessment degrees after all criteria in $G_{I(l)}$ subset are assessed. $m_{q,I(l)}(x_i)$ and $m_{H,I(l)}(x_i)$ can be obtained combining basic probable mass of $m_{q,j}(x_i)$ and $m_{H,j}(x_i)$ for all $q = 1,2,\ldots,p$ and $j = 1,2,\ldots,l$.

According to the definitions and subjects mentioned above, evidence-based reasoning recursive algorithm can be summarized as follows:

$$\{H\}: m_{q,I(l+1)}(x_i) = K_{I(l+1)} \left[ m_{q,I(l)}(x_i)m_{q,j+1}(x_i) + m_{H,I(l)}(x_i)m_{q,I(j+1)}(x_i) + m_{H,I(l)}(x_i)m_{H,I(j+1)}(x_i) \right].$$

$$m_{H,I(l+1)}(x_i) = \tilde{m}_{H,I(l+1)}(x_i) + \tilde{m}_{H,I(l+1)}(x_i),$$

$$\{H\}: \tilde{m}_{H,I(l+1)}(x_i) = K_{I(l+1)} \left[ \tilde{m}_{H,I(l)}(x_i)\tilde{m}_{H,I(l+1)}(x_i) \right].$$

$$\{H\}: \tilde{m}_{H,I(l+1)}(x_i) = K_{I(l+1)} \left[ \tilde{m}_{H,I(l)}(x_i)\tilde{m}_{H,I+1(l)}(x_i) + \tilde{m}_{H,I(l)}(x_i)\tilde{m}_{H,J(l)}(x_i) + \tilde{m}_{H,I(l)}(x_i)\tilde{m}_{H,J(l+1)}(x_i) \right].$$

$$K_{I(l+1)} = \left[ 1 - \sum_{u=1}^{p} \sum_{j=1}^{l} m_{u,I(l)}(x_i)m_{f,I+1(l)}(x_i) \right]^{-1}$$

$K_{I(l+1)}$ is a normalization factor through which $\sum_{q=1}^{p} m_{q,I(l+1)}(x_i) + m_{H,I(l+1)}(x_i)=1$. Note that $m_{q,I(l)}(x_i)=m_{q,1}(x_i) (q = 1, 2, \ldots, p)$ and $m_{H,I(l)}(x_i)=m_{H,1}(x_i)$. Besides, it should be considered that the criteria existing in $G$ are numbered randomly, that is the results of $m_{q,I(l)}(x_i), (q = 1, 2, \ldots, p)$, and $m_{H,I(l)}(x_i)$ are independent of the sum order of criteria [47].
In the evidence-based reasoning approach, after that all \( n \) criteria are aggregated, the combined belief degree \( \beta_q \) is directly calculated from the below equation:

\[
\{H_q\} : \beta_q(x_i) = \frac{m_{q,J(n)}(x_i)}{1-m_{H,J(n)}(x_i)}.
\]

\[
\{H\} : \beta_H(x_i) = \frac{\bar{m}_{H,J(n)}(x_i)}{1-\bar{m}_{H,J(n)}(x_i)}.
\]

\( \beta_H \) is a belief degree that is not devoted to any assessment degree after all \( n \) criteria are appraised. Indeed, this parameter defines the violation degree existing in the assessment procedure. Therefore, we will have: \( \sum_{q=1}^{p} \beta_q(x_i) + \beta_H(x_i) = 1. \)

Ultimately, utility functions that actually indicate the relative importance of possible exchange rate values based on an aggregation of experts' opinions are indicated in the form of \([u_{min}, u_{max}]\), and are indicated by \( u(H_a) \) and calculated as follows:

\[
u_{\text{max}}(x) = \sum_{q=1}^{p-1} B_q u(H_q) + (B_q + B_H) u(H_1) \]

\[
u_{\text{min}}(x) = (B_N + B_H) u(H_N) + \sum_{q=2}^{p} B_q u(H_q) \]

\[
u_{\text{Avg}}(y) = \frac{\nu_{\text{max}}(y) + \nu_{\text{min}}(y)}{2}
\]

Where, \( 1 \leq q \leq p \) and \( H_N \) have the minimum preference degree and \( H_1 \) has the maximum preference degree. Besides, \( B_N \) is allocated the maximum belief degree, and \( B_1 \) is allocated the minimum belief degree. Notice that if the main assessment of \( S(a_j(x_i)) \) is complete, we will have \( B_H = 0 \) [48].

4. Description of portfolio optimization approach

Portfolio optimization is the process of selecting the best portfolio (asset distribution) among the set of all portfolios being considered, according to the specific objective. The risk level and yield of invested assets are two essential parameters in deciding on an investment. Most investors want maximize their efficiency at a specific level of risk or minimize risk for a certain efficiency level. The individuals invest according to their expected utility and in the hope of more gaining in the future, ignore today's consumption. The utility function of each investor is ascertained according to that individual's preferences [49]. Proposing the mean-variance model, Markowitz indicated that selecting a portfolio makes it possible to reduce investment risk at a certain efficiency level. This possibility is attributed to a lack of thorough correlation between the efficiencies of different stocks. Optimal portfolio selection is often carried out by exchange between risk and yield such that the more investors' expected efficiency, the more risk of a portfolio. Accordingly, identifying the efficient frontier pertinent to the portfolio enables
investors to obtain their maximum expected efficiency according to the utility function and risk-aversion and risk-taking. Our main model based on the primary framework submitted by Markowitz is developed. On the other hand, as presented by Markowitz, the classical portfolio selection problem completely disregards the uncertainty of the expected asset’s returns and the covariance matrix of assets’ returns. It is assumed that these parameters are capable of representing the inherent uncertainty associated with the investment returns. Actually, as these parameters are, most of the times, calculated from past data, they are themselves subject to uncertainty. Not acknowledging the uncertainty in the models’ parameters substantially degrades the performance of the optimal solution calculated using these Models. The robust formulation of an optimization problem considers the nominal values of the uncertain parameters and the deviations from these nominal values.

4.1. Conventional portfolio model

In 1952, proposing the portfolio optimization problem-solving model (mean-variance theory) for the first time, Markowitz [14] expressed the issue as a Quadratic Programming to minimize portfolio risk when the expected efficiency is a constant value. The main assumption of this model is that all investors are risk-avert. This problem consisted of two applied constraints, according to which the sum of stock weights must be equal to one, and also the weight of each stock in the selected portfolio must be a non-negative number. The standard form of the mean-variance model is as follows:

\[
\text{Max } \mu_p = \sum_{i=1}^{n} x_i \mu_i
\]  \hspace{1cm} (18)

Subject to:

\[
\sum_{i=1}^{n} \sum_{j=1}^{n} \sigma_{ij} x_i x_j = \sigma_p^2
\]  \hspace{1cm} (19)

\[
\sum_{i=1}^{n} x_i = 1
\]  \hspace{1cm} (20)

\[
x_i \geq 0, \quad (i = 1, \ldots, n)
\]  \hspace{1cm} (21)

\[
x_j \geq 0, \quad (j = 1, \ldots, n)
\]  \hspace{1cm} (22)

In the Markowitz model, the primary point is that the general risk criterion is variance or standard deviation. This criterion that is generally distributed for stock and exchanged in an efficient market is acceptable. Now, suppose these two features do not exist for a store, for any reason, including ambiguities or severe and unpredictable fluctuations in the influential factors to the stock market. In that case, the variance will not be a suitable index for stock risk. Therefore, another solution must be found.

4.2. Robust optimization

Over the past few years, due to the ever-increasing attention paid to taking account of uncertainty in decision-making processes, many investigations have been conducted in this field, and various approaches have been employed. One of the most efficient methods used in this field
is RO [50]. RO is one of the approaches employed to deal with uncertainty in optimization problems and aims to model optimization problems with uncertain data to make the obtained results suitable for all (or most) of the imaginable cases of uncertain parameters, Cornuejols and Tütüncü [51]. A robust answer consists of robustness in terms of optimality and robustness in terms of feasibility. Robustness in terms of feasibility means that the desired response for all (or most) of the imaginable cases remains feasible for uncertain parameters. Robustness in terms of optimality implies that the preferred solution for all (or most) potential cases for uncertain parameters remains close to optimal value or has the minimum standard deviation from the optimal value, Pishvaee et al.[52].

Soyster [53] was the first one investigating the linear optimization problem with uncertain datasets. He generated extremely cautious answers on the basis of the assumption that uncertain datasets are within a certain range. Afterward, Mulvey et al. [54] and Yu and Li [55] proposed scenario-based robust models. El Ghaoui et al. [56] proposed a robust counterpart programming for semidefinite linear and quadratic problems based on the ellipsoidal uncertainty set assumption. Ben-tal and Nemirovsk [57] i indicated that if datasets of a linear optimization problem with uncertainty belong to a semi-ellipsoidal set (conic, quadratic, semidefinite), its robust counterpart or an approximation of it can be solved in polynomial time. Ben-tal and Nemirovski [58] investigated linear programming problems with inaccurate datasets and enumerated the features of an uncertain convex set for the robust counterpart to be exactly equivalent to the initial problem. Ben-tal and Nemirovski [57, 58] indicated that the robust counterpart is conic and quadratic when the uncertainty set is ellipsoidal. Ben-tal and Nemirovski [59] investigated the impact of uncertainty on optimization problems. Considering the semi-ellipsoidal set as the uncertainty set, they proposed a robust counterpart for linear optimization and a relationship for the probability of the problem constraints violation due to uncertainty.

Bertsimas and Sim [60] investigated optimization problems for discrete problems and proposed a robust counterpart model that is computable and capable of being adjusted in terms of robustness, and examined the probability of constraints' violation. Bertsimas and Sim [28] proposed a robust linear counterpart for linear programming problems with an uncertain dataset, which can be adjusted in terms of cautiousness and balanced between acceptability and optimality. They also developed the proposed model for discrete programming problems and examined the probability of constraints' violation. Now, given that the robustification approach employed in this investigation is of Bertsimas and Sim type, this model will be explained in the following. In order to examine this method, the general form of the linear optimization problem can be considered as follows:

$$\text{Max } \sum_{i} c_i x_i$$  \hspace{1cm} (23)

$$\sum_{i} \tilde{a}_{ij} x_i \leq b_j, \forall j$$  \hspace{1cm} (24)

$$x_i \geq 0, \forall i$$  \hspace{1cm} (25)

$x_i$ is the decision variable, $j$ is the number of constraints. $c_i$ and $b_j$ have certain values, while $\tilde{a}_{ij}$ is uncertain. Uncertainty of $\tilde{a}_{ij}$ is defined as follows:
\[ \hat{a}_{ij} = a_{ij} + \xi_{ij} \hat{a}_{ij} \]  

where \( a_{ij} \) is the nominal value of data, \( \hat{a}_{ij} \) is a positive constant uncertainty value that indicates disruption value, \( \xi_{ij} \hat{a}_{ij} \) is a random value within \([-1,1]\) range and is pertinent to uncertainty, which is indicated as a set.

Therefore, using multi-dimensional uncertainty set and duality theorem, Bertsimas and Sim formulated a robust counterpart of linear optimization problems, which can be adjusted in terms of cautiousness and balanced between feasibility and optimality. This formulation can be employed for discrete optimization problems. Robust counterpart formulation of equations (23) to (25) using this approach are as follows, Bertsimas and Sim [28]:

\[
\text{Max } \sum_{i} c_{i} x_{i} \quad (27)
\]

Subjected to:

\[
\sum_{i} a_{ij} x_{i} + \sum_{i \in J_j} p_{ij} \leq b_{j}, \quad \forall j \quad (28)
\]

\[
z_{j} + p_{ij} \geq \hat{a}_{ij} y_{i}, \quad \forall j, i \in J_j \quad (29)
\]

\[
-y_{i} \leq x_{i} \leq y_{i}, \quad \forall i \quad (30)
\]

\[
p_{ij} \geq 0, \quad \forall j, i \in J_j \quad (31)
\]

\[
y_{i} \geq 0, \quad \forall i \quad (32)
\]

\[
z_{j} \geq 0, \quad \forall j \quad (33)
\]

The cautiousness of each constraint can be determined by \( \Gamma_{j} \), whose maximum value is equal to the number of \( \xi_{i} \neq 0 \) value or \( |J_j| \).

After pointing out the necessary subjects, we will address our model and explain it thoroughly in the following section.

5. Case study

In order to analyze any real issue with a mathematical modelling tool, a combination of assumptions with regard to the behavior of system must be considered. In this regard, the return fluctuates within a symmetric range, and all investors have an identical single-period time horizon. Moreover, personal incomes are tax-free, that is investors do not distinction between dividend profit and capital profit, and trade in the market is costless. Furthermore, this problem is not affected by inflation. Moreover, the stock price cannot be solely impacted by any capital according to sell and buy decisions. Ultimately, investors prefer a higher efficiency level at a certain level of risk, and for a certain level of efficiency seek a minimum level of risk.
5-1 symbols of the problem

Parameters

- $R_i$: Expected efficiency
- $\beta_i$: Measurement unit of systematic risk
- $R_m$: Index efficiency rate
- $e_i$: Random error
- $\hat{r}_i$: The expected value of asset efficiency
- $\hat{a}_i$: Maximum possible deviations
- $\Gamma$: Robustness cost
- $n$: The number of stocks
- $t$: A unique upper bound for all stocks

Variables:

- $x_i$: The weight value of the stock $i$ if selected

Thereby, we have:

$$\text{Max} R_p = \sum_{i=1}^{n} \bar{r}_i x_i$$  \hspace{1cm} (34)

Subjected to:

$$\sum_{i=1}^{n} x_i = 1$$  \hspace{1cm} (35)

$$0 \leq x_i \leq U$$  \hspace{1cm} (36)

$x_i$ is the variable of $i$ decision, and $\bar{r}_i$ is uncertain. Additionally, constraint (1) ascertains that the whole budget must be devoted to different alternatives, it is not allowed to use additional budget. Constraint (2) or sign constraint expresses that using short sell is impossible.

Furthermore, in a linear programming problem in the form mentioned above, the uniform distribution assumption for efficiency can be considered as follows, according to the stated assumption and ignoring the correlation between stocks:

$$\bar{r}_i \sim U (\bar{r}_i - \hat{a}_i, \bar{r}_i + \hat{a}_i)$$  \hspace{1cm} (37)

Ultimately, robust counterpart formulation of equations (34) to (36) using Bertsimas and Sim approach are as follows:
Max $U$ \hspace{1cm} (38)

**Subjected to:**

$$U - \sum_{i=1}^{n} \tilde{r}_i x_i + z \Gamma + \sum_{i=1}^{n} P_i \leq 0 \quad (39)$$

$$z + P_i \geq \tilde{a}_i y_i \quad \forall i = 1, \ldots, n \quad (40)$$

$$-y_i \leq x_i \leq y_i \quad \forall i = 1, \ldots, n \quad (41)$$

$$\sum_{i=1}^{n} x_i = 1 \quad \forall i = 1, \ldots, n \quad (42)$$

$$x_i \leq t \quad \forall i = 1, \ldots, n \quad (43)$$

$$x_i, y_i, P_i \geq 0 \quad \forall i = 1, \ldots, n \quad (44)$$

In the proposed formulation, as mentioned earlier, it is assumed that the uncertain coefficient $\tilde{a}_i$ is a uniform random value within the range of $[\tilde{r}_i - \tilde{a}_i, \tilde{r}_i + \tilde{a}_i]$. In this case, if the number of coefficients with uncertainty, that is $\Gamma_i$, changes in the raw $i^{th}$ in which $1 \leq i \leq n$, the answer will still be feasible, and if more than number of coefficients change, the constraint will probably be rejected. In addition, the parameter $\Gamma_i$ might not be integer and can have a value within range of $[0, |\Gamma_i|]$.

6. **Computational results**

In this investigation, the collected data belong to 10 firms active in Tehran capital market between 2019, April to 2021, July. The following constraint is applied for the selection of stock sets: A) In the study period, they do not face trading halt for more than half of the year B) Their financial year ends on 20 March C) They have the lowest risk compared to the rest, and their monthly returns are not less than 10%. D) They have clear and complete financial statements and information.

As discussed earlier, the correlation coefficient between dollar rate and each stock's prices plays a pivotal role in the success or failure of the investment, particularly in our country's stock market, and therefore, has captured more interest nowadays than before. Analyzing the correlation coefficient between stock price and stock exchange index, researchers employed the Beta coefficient in order to appraise the rank of stocks. In the present study, due to the impact of dollar rate on the stock market index and the rest of Iran financial markets, the dollar exchange rate is employed as a pivotal factor. As a result, in this section, the quantity of the dollar exchange rate which is an uncertain parameter will be computed by using DST. Afterward, based on the robust Bertsimas and Sim model, the result of solving the model and optimal values of the stock portfolio for different values of $\Gamma$ will be presented. Accordingly, experts ascertain the belief degree about the dollar exchange rate, depicted in Table 1.

(Insert Table 1 almost here)
It is seen that, the rate of dollar exchange which includes (Scenario A, Scenario B, Scenario C, Scenario D) with values (150000, 200000, 250000, 300000) is predicted using DTS in the [19.607, 25.500] range. Therefore, its percentage of deviation from the mean is in the [-0.130, 0.130] range. Consequently, considering the value of β between stock returns and dollar returns and utilizing this information, the uncertain amount of stock returns is computed intermittently to use the Bertsimas and Sim robust optimization approach to model the stock portfolio represented in Table 2.

(Insert Table 2 almost here)

Finally, due to the linearity of the achieved model, it is solved by GAMS 24.1.3 on a personal computer based on 2.8 GHz at 0.016 seconds with the CPLEX solver. The results are shown in Table 3.

(Insert Table 3 almost here)

Solving the model with the software Gams, we formed an optimal stock portfolio. It is worth mentioning that by increasing the value of Γ due to adding the number of parameters with uncertainty, the return of the optimal portfolio of stocks decreases. If Γ=0, the model is completely certain, and if Γ=10, the relevant parameters can indeed have the maximum fluctuation from central limitation. In other words, when the model is deterministic, due to the higher average returns relative to the other stocks, stock 1, stock 4, stock 6, and stock 8 are selected with the optimal portfolio return of 0.194. Accordingly, the return of the optimal portfolio remarkably declines to 0.104 when uncertainty embraces all of the parameters. The logic of this model is that the probability that all uncertain coefficients have their worst value in nature is less, and usually, some of them face variation. The results are represented in Figure 2.

(Insert Fig.2 almost here)

7. Conclusions

An inherent feature in human mental judgments that should be given special attention to in decisions is Uncertainty. In this study, a novel approach for the portfolio construction problem is proposed in order to deal with data uncertainty, increasing conservatism levels of the investment process, decreasing computational complexity, and assessing comprehensive of stocks. Accordingly, the present study presents a stock portfolio optimization model which is considering the dollar exchange rate, and seeks to consider information deficiencies to improve performance using the logic based on DTS and RO. It is also formulated in a given atmosphere. Finally, in this paper, in order to demonstrate the applicability of the submitted model and exhibit the efficacy and effectiveness of the presented method, a real-life case study from the Tehran stock exchange is implemented. Though this model is implemented here for Iran Stock Market, it can be applied for constructing portfolios in any Stock Exchanges around the world but selection of critical factors can vary in different stock exchanges. For future studies, to enhance the robustness of the model researchers can think of hybridizing DS evidence theory with other uncertainty handling tools like soft sets and rough sets. Moreover,
data-driven robust optimization (DDRO) approach can be employed for proposing data-driven robust portfolio optimization (DDRPO) models (see [61], [62],[63], [64]).

Biographies

Amirhossein Eskorouchi was born on June 19, 1996 in Tehran, Iran. He graduated Bachelor of Science in Industrial Engineering from the University of Garmsar, Semnan, Iran, and Master of Science from Iran University of Science and Technology, Tehran, Iran.

Emran Mohammadi is working as an Associate Professor at the Department of Industrial Engineering at the Iran University of Science and Technology, Tehran, Iran. More than 50 students have completed their respective degrees under his supervision. He has published over 100 research papers in reputed international journals, with over 700 citations. Several national and international awards have been conferred on him. He is also a member of several editorial boards in well reputed international journals. He is a member of various scientific committees in the country.

Seyed Jafar Sajadi is working as a professor at the Iran University of Science and Technology, Tehran, Iran. He has published more than 400 research articles in well reputed international journals, with over 4400 citations. Additionally, several national and international awards have been conferred on him. He is also a member of several editorial boards in well reputed international journals.

Figure captions

Fig 1. Conceptual model of this research

Page 4

Fig 2. Portfolio Returns Based on Different Levels of Conservatism

Page 14

Table captions

Table 1. Belief Degree Based on Evidence Theory

Page 13

Table 2. The bound of returns based on sharp multi-factor model

Page 13

Table 3. The optimal answer based on different levels of conservatism

Page 14
Determining investment alternatives

Determining the effective factors on stock return

Collecting the required basic data

Fundamental analysis of investment alternatives based on financial ratios

Prioritizing effective factors

Detecting behavioral scenarios in terms of effective factors changes

Get expert opinion on each scenario

Combining experts opinions according to evidence theory (Dempster-Shafer)

Determining the bound of uncertainty for the effective factors

Identification stocks with a good foundation according to data

Filtering based on thresholds of maximum expected return and minimum expected risk

Determining how stock prices are affected by effective factors based on sharp multi-factor model

Determining the boundaries of stock return changes

Robust portfolio optimization

Model robustness analysis based on different levels of conservatism

Fig 1.
**Fig 2.**

**Table 1.**

<table>
<thead>
<tr>
<th>DM</th>
<th>Weight</th>
<th>Scenario A</th>
<th>Scenario B</th>
<th>Scenario C</th>
<th>Scenario D</th>
<th>$B_H$</th>
</tr>
</thead>
<tbody>
<tr>
<td>A1</td>
<td>0.100</td>
<td>0.206</td>
<td>0.086</td>
<td>0.107</td>
<td>0.206</td>
<td>0.392</td>
</tr>
<tr>
<td>A2</td>
<td>0.100</td>
<td>0.000</td>
<td>0.050</td>
<td>0.300</td>
<td>0.600</td>
<td>0.050</td>
</tr>
<tr>
<td>A3</td>
<td>0.450</td>
<td>0.100</td>
<td>0.000</td>
<td>0.100</td>
<td>0.300</td>
<td>0.600</td>
</tr>
<tr>
<td>A4</td>
<td>0.350</td>
<td>0.400</td>
<td>0.150</td>
<td>0.050</td>
<td>0.000</td>
<td>0.400</td>
</tr>
</tbody>
</table>

**Table 2.**

<table>
<thead>
<tr>
<th></th>
<th>Stock1</th>
<th>Stock2</th>
<th>Stock3</th>
<th>Stock4</th>
<th>Stock5</th>
<th>Stock6</th>
<th>Stock7</th>
<th>Stock8</th>
<th>Stock9</th>
<th>Stock10</th>
<th>Upper return</th>
<th>Average return</th>
<th>Tolerance</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.286</td>
<td>0.241</td>
<td>0.223</td>
<td>0.262</td>
<td>0.189</td>
<td>0.255</td>
<td>0.170</td>
<td>0.224</td>
<td>0.164</td>
<td>0.173</td>
<td>0.112</td>
<td>0.199</td>
<td>0.087</td>
</tr>
<tr>
<td></td>
<td>0.112</td>
<td>0.016</td>
<td>0.065</td>
<td>0.121</td>
<td>0.063</td>
<td>0.138</td>
<td>0.063</td>
<td>0.120</td>
<td>0.062</td>
<td>0.091</td>
<td>0.129</td>
<td>0.144</td>
<td>0.082</td>
</tr>
<tr>
<td></td>
<td>0.199</td>
<td>0.129</td>
<td>0.144</td>
<td>0.192</td>
<td>0.126</td>
<td>0.197</td>
<td>0.117</td>
<td>0.172</td>
<td>0.113</td>
<td>0.132</td>
<td>0.079</td>
<td>0.050</td>
<td>0.070</td>
</tr>
<tr>
<td></td>
<td>0.262</td>
<td>0.231</td>
<td>0.169</td>
<td>0.190</td>
<td>0.230</td>
<td>0.257</td>
<td>0.257</td>
<td>0.000</td>
<td>0.000</td>
<td>0.164</td>
<td>0.079</td>
<td>0.050</td>
<td>0.062</td>
</tr>
<tr>
<td></td>
<td>0.086</td>
<td>0.090</td>
<td>0.106</td>
<td>0.169</td>
<td>0.300</td>
<td>0.020</td>
<td>0.000</td>
<td>0.025</td>
<td>0.000</td>
<td>0.013</td>
<td>0.091</td>
<td>0.090</td>
<td>0.017</td>
</tr>
<tr>
<td></td>
<td>0.012</td>
<td>0.013</td>
<td>0.013</td>
<td>0.017</td>
<td>0.300</td>
<td>0.000</td>
<td>0.000</td>
<td>0.017</td>
<td>0.021</td>
<td>0.011</td>
<td>0.011</td>
<td>0.011</td>
<td>0.010</td>
</tr>
<tr>
<td></td>
<td>0.010</td>
<td>0.011</td>
<td>0.011</td>
<td>0.014</td>
<td>0.300</td>
<td>0.016</td>
<td>0.000</td>
<td>0.017</td>
<td>0.021</td>
<td>0.011</td>
<td>0.011</td>
<td>0.011</td>
<td>0.010</td>
</tr>
<tr>
<td></td>
<td>0.005</td>
<td>0.057</td>
<td>0.060</td>
<td>0.067</td>
<td>0.075</td>
<td>0.300</td>
<td>0.088</td>
<td>0.091</td>
<td>0.093</td>
<td>0.114</td>
<td>0.067</td>
<td>0.075</td>
<td>0.057</td>
</tr>
<tr>
<td></td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.300</td>
<td>0.000</td>
<td>0.300</td>
<td>0.000</td>
<td>0.014</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.300</td>
<td>0.000</td>
<td>0.300</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
</tbody>
</table>

**Table 3.**

<table>
<thead>
<tr>
<th>Γ</th>
<th>Stock 1</th>
<th>Stock 2</th>
<th>Stock 3</th>
<th>Stock 4</th>
<th>Stock 5</th>
<th>Stock 6</th>
<th>Stock 7</th>
<th>Stock 8</th>
<th>Stock 9</th>
<th>Stock 10</th>
<th>Portfolio returns</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0.300</td>
<td>0.000</td>
<td>0.000</td>
<td>0.300</td>
<td>0.000</td>
<td>0.300</td>
<td>0.000</td>
<td>0.100</td>
<td>0.000</td>
<td>0.000</td>
<td>0.194</td>
</tr>
<tr>
<td>1</td>
<td>0.200</td>
<td>0.000</td>
<td>0.000</td>
<td>0.247</td>
<td>0.000</td>
<td>0.300</td>
<td>0.000</td>
<td>0.253</td>
<td>0.000</td>
<td>0.000</td>
<td>0.173</td>
</tr>
<tr>
<td>2</td>
<td>0.187</td>
<td>0.000</td>
<td>0.000</td>
<td>0.231</td>
<td>0.000</td>
<td>0.281</td>
<td>0.000</td>
<td>0.300</td>
<td>0.000</td>
<td>0.000</td>
<td>0.156</td>
</tr>
<tr>
<td>3</td>
<td>0.154</td>
<td>0.000</td>
<td>0.169</td>
<td>0.190</td>
<td>0.000</td>
<td>0.230</td>
<td>0.000</td>
<td>0.257</td>
<td>0.000</td>
<td>0.000</td>
<td>0.141</td>
</tr>
<tr>
<td>4</td>
<td>0.086</td>
<td>0.090</td>
<td>0.094</td>
<td>0.106</td>
<td>0.000</td>
<td>0.300</td>
<td>0.000</td>
<td>0.144</td>
<td>0.000</td>
<td>0.000</td>
<td>0.131</td>
</tr>
<tr>
<td>5</td>
<td>0.012</td>
<td>0.013</td>
<td>0.013</td>
<td>0.017</td>
<td>0.300</td>
<td>0.020</td>
<td>0.300</td>
<td>0.000</td>
<td>0.000</td>
<td>0.025</td>
<td>0.126</td>
</tr>
<tr>
<td>6</td>
<td>0.010</td>
<td>0.011</td>
<td>0.011</td>
<td>0.014</td>
<td>0.300</td>
<td>0.016</td>
<td>0.300</td>
<td>0.017</td>
<td>0.021</td>
<td>0.011</td>
<td>0.116</td>
</tr>
<tr>
<td>7</td>
<td>0.054</td>
<td>0.057</td>
<td>0.060</td>
<td>0.067</td>
<td>0.075</td>
<td>0.300</td>
<td>0.088</td>
<td>0.091</td>
<td>0.093</td>
<td>0.114</td>
<td>0.109</td>
</tr>
<tr>
<td>8</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.100</td>
<td>0.000</td>
<td>0.300</td>
<td>0.000</td>
<td>0.300</td>
<td>0.000</td>
<td>0.300</td>
<td>0.104</td>
</tr>
<tr>
<td>9</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.100</td>
<td>0.000</td>
<td>0.300</td>
<td>0.000</td>
<td>0.300</td>
<td>0.000</td>
<td>0.300</td>
<td>0.104</td>
</tr>
<tr>
<td>10</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.100</td>
<td>0.000</td>
<td>0.300</td>
<td>0.000</td>
<td>0.300</td>
<td>0.000</td>
<td>0.300</td>
<td>0.104</td>
</tr>
</tbody>
</table>